

Sıralı Küme Örneklemesi ile Kumaraswamy Dağılımı Parametrelerinin Tahmin Edilmesinde Genetik Algoritma Kullanılması

Adil KILIÇ*¹, Güvenç ARSLAN²

^{1,2}Kırıkkale Üniversitesi, Fen Edebiyat Fakültesi, İstatistik Bölümü, 71450, Kırıkkale, Türkiye

(Alınış / Received: 17.10.2018, Kabul / Accepted: 05.04.2019, Online Yayınlanma / Published Online: 30.08.2019)

Anahtar Kelimeler

Kumaraswamy dağılımı,
Sıralı küme örnekleme,
En çok olabilirlik tahmini,
Genetik algoritma,
R yazılımı

Özet: Bu çalışmada, Kumaraswamy dağılımının parametrelerinin en çok olabilirlik yöntemi ile tahmin edilmesi genetik algoritma yaklaşımı kullanılarak araştırılmıştır. Ayrıca basit rasgele örnekleme göre daha iyi sonuç verebileceği düşünülerek parametrelerin tahmin edilmesinde sıralı küme örnekleme de incelenmiştir. Genetik algoritma yaklaşımı, Kumaraswamy dağılımı parametrelerinin pozitif olma koşulunun hesaba katılması nedeniyle tercih edilmiştir. Ek olarak genetik algoritma yaklaşımında en çok olabilirlik fonksiyonunun türev bilgisine ihtiyaç duyulmaması da hesaplamalarda kolaylık sağlamaktadır. Genetik algoritma kullanılarak elde edilen her iki örnekleme yöntemine ait olabilirlik tahmin edicilerinin performanslarının karşılaştırılması için yan, hata kareler ortalaması ve etkinlikleri hesaplanmıştır. Simülasyon çalışmasındaki hesaplamalar için R yazılımı ve ilgili paketler kullanılmıştır.

On Estimating Parameters of the Kumaraswamy Distribution with Ranked Set Sampling Using Genetic Algorithms

Keywords

Kumaraswamy distribution,
Ranked set sampling,
Maximum likelihood
estimation,
Genetic algorithm,
R software

Abstract: In this paper, genetic algorithm approach is used to estimate parameters of the Kumaraswamy distribution with maximum likelihood method. In addition ranked set sampling is used since it is expected to give better results in comparison to simple random sampling. Genetic algorithm approach is chosen because it is relatively more convenient in terms of satisfying positivity constraints for the parameters of the Kumaraswamy distribution. Also there is no need to use derivatives in the genetic algorithm approach. Bias, MSE and efficiency is calculated to compare performances of maximum likelihood estimators for ranked set sampling and simple random sampling obtained by using genetic algorithms. The R software and related packages are preferred for calculations in the simulation study.

1. Giriş

Kumaraswamy dağılımı 1980 yılında hidroloji uygulamalarında alttan ve üstten sınırlı değişkenler için P. KUMARASWAMY tarafından önerilmiştir [1]. [0, 1] kapalı aralığında sürekli olasılık dağılımlar ailesinden iki sınırlı bir dağılımdır. Beta dağılımına benzerlik göstermesinin yanında olasılık yoğunluk fonksiyonu ve birikimli dağılım fonksiyonlarının kapalı formda yazılabilmesinden dolayı özellikle simülasyon çalışmalarında daha kolay bir kullanımı mevcuttur [2]. Bu dağılıma ait olasılık yoğunluk fonksiyonu ve birikimli dağılım fonksiyonu sırasıyla (1) ve (2) eşitliklerindeki gibidir.

$$f(x; a, b) = abx^{a-1}(1-x)^{b-1} \quad (1)$$

$$F(x; a, b) = 1 - (1-x^a)^b \quad (2)$$

Burada $0 < x < 1$ ve $a, b > 0$ olmak üzere a ve b dağılımın şekil parametreleridir. M. A. Hussian bu dağılımın hem basit rasgele hem de sıralı küme örnekleme ile en çok olabilirlik ve bayes tahmin edicilerini *MATHEMATICA* yazılımını kullanarak hesaplamış ve karşılaştırmıştır [3]. Bu çalışmada ise bu iki örnekleme yöntemi ile elde edilen en çok olabilirlik tahmin edicilerinin hesaplanmasında genetik algoritma yaklaşımı kullanılmıştır.

2. Sıralı Küme Örneklemesi

Kitle hakkında çıkarım yapabilmek için o kitleden, kitleyi iyi temsil edebilecek örneklem oluşturulması

gerekir. Dolayısıyla kitleden seçilen örneklemin hangi yöntemle seçileceği kitle hakkında doğru ve iyi bir çıkarımda bulunabilmek için oldukça önemlidir. Günümüze kadar birçok farklı örnekleme yöntemi önerilmiştir. Örnekleme yöntemleri içinde yaygın olarak bilinen yöntem basit rasgele örnekleme yöntemidir. Bu yöntemde kitledeki her birim eşit seçilme şansına sahiptir. X_1, X_2, \dots, X_n basit rasgele örnekleme yöntemi ile elde edilmiş aynı $F(x)$ dağılımlı rasgele değişkenler koleksiyonu olmak üzere bu rasgele değişkenler birbirinden bağımsızdır. Örneklem seçilecek kitle eğer çok büyük veya geniş bir bölgede düzensiz bir biçimde yayılmışsa bu yöntem ile örneklem oluşturmak sakıncalı olabilir.

Bu çalışmada basit rasgele örnekleme alternatif olarak sıralı küme örnekleme de incelenmiştir. Sıralı küme örnekleme ilk olarak 1952 yılında McIntyre tarafından Avustralya’da mera verimini tahmin etmek için kullanılmıştır [4]. Bu örnekleme yöntemi her bir gözlem için tam ölçüm yapılmadan örnekleme seçilecek birimlerin belirlenmesinin sıralamaya dayalı olduğu bir yöntemdir. Bu yöntemde örnekleme alınacak olan birimler ilgili değişkene göre sıralanır. Daha sonra sıralama gözleme dayalı olarak ya da yardımcı değişken kullanılarak yapılır. Sıralama işlemi yapıldıktan sonra bu sıralamaya göre tam ölçümü yapılacak olan gözlemler alınıp tam olarak ölçülür. Burada sıralama hatası söz konusu olabilir ancak yine de eşit örneklem çaplı basit rasgele örnekleme göre daha iyi tahminler elde edilebilir. Tahminlerdeki duyarlılık örneklem çapına ve sıralamadaki hatalar gibi faktörlere bağlı olarak değişiklik gösterebilir [5].

Takahasi ve Wakimoto sıralı küme örneklemesinin ortalama tahmincisinin basit rasgele örnekleminin kitle tahmincisine kıyasla daha küçük varyanslı ve kitle ortalaması için yansız olduğunu göstermiştir [6]. Dell ve Clutter sıralama hatası olsa da olmasa da SKÖ ortalamasını yine de yansız olduğunu göstermiştir [7]. Stokes SKÖ için sıralamada görsel değerlendirme veya uzman görüşü yerine birimleri sıralamak için yardımcı değişkenler kullanılmasını önermiştir [8]. Kitle ortalama tahmininin doğruluğunu arttırmak için Samawi tarafından Tabakalandırılmış Sıralı Küme Örneklemesi önerilmiştir [9]. Samawi ve arkadaşları dağılımın asimetrik olduğu durumda kitle tahmincisini tanımlamak için Extreme Sıralı Küme Örneklemesini kullanmıştır [10]. Al-Saleh ve Al-Kadiri Double Sıralı Küme Örneklemesi geliştirmiştir [11].

Al-Saleh ve Al-Omari SKÖ’ nün genelleştirilmiş hali olarak Çok Aşamalı SKÖ metodunu sunmuştur [12]. Mutlak asimetrik dağılımlarda daha iyi sonuç veren Quartile SKÖ’ yi önermiştir [13].

Adım adım sıralı küme örnekleme ile örneklem seçme işlemi aşağıdaki gibi tanımlanır.

Adım 1. İlgilenilen kitleden rasgele olarak m^2 tane örneklem seçilir.

Adım 2. Seçilen m^2 çaplı bu örneklem rasgele olarak m adet m çaplı kümeye dağıtılır.

Adım 3. Bu m çaplı kümelerin her biri küçükten büyüğe doğru tam ölçmeden görsel yolla veya kolay ve pratik bir ölçümle sıralanır.

Adım 4. Birinci kümeden en küçük birim, ikinci kümeden ikinci sıradaki birim ve aynı sıra takip edilerek m . kümeden m . sıradaki yani en büyük olan birim seçilir.

Adım 5. Bu prosedür ile seçilen m çaplı örneklem istenen hassasiyetle ölçümü yapılır ve daha büyük bir sıralı küme örnekleme için aynı adımlar tekrarlanır.

Bu adımlar gözlem sayısı $m=3$ ve döngü sayısı $r=2$ için 6 birimlik bir sıralı küme örnekleminin seçimi Şekil 1 de şematize edilmiştir. m , kümelerdeki birim sayısını ve r de döngü sayısını göstermektedir. $X_{(i:m)r}$: r inci döngüdeki m birimlik örnekleme i inci sıradaki değişkeni temsil etmektedir.

döngü / sıra	1.	2.	3.
$r = 1$	$X_{(1:3)1}$	$X_{(2:3)1}$	$X_{(3:3)1}$
	$X_{(1:3)1}$	$X_{(2:3)1}$	$X_{(3:3)1}$
	$X_{(1:3)1}$	$X_{(2:3)1}$	$X_{(3:3)1}$
$r = 2$	$X_{(1:3)2}$	$X_{(2:3)2}$	$X_{(3:3)2}$
	$X_{(1:3)2}$	$X_{(2:3)2}$	$X_{(3:3)2}$
	$X_{(1:3)2}$	$X_{(2:3)2}$	$X_{(3:3)2}$

Şekil 1. SKÖ Seçim Şeması

Bu örnekleme yönteminde her m birimlik alt kümeler birbirinden bağımsız olduğu için kümelerin aynı sıradaki değişkenlerin gösterimi aynı notasyonla yapılmış olsa da birbirinden farklıdır.

3. Parametre Tahmini

Burada basit rasgele ve sıralı küme örnekleme için Kumaraswamy dağılımı parametrelerine ait en çok olabilirlik tahmin edicilerinin matematiksel olarak elde edilmesi gösterilmiştir.

3.1. BRÖ için en çok olabilirlik tahmini

X_1, X_2, \dots, X_n , n boyutlu a ve b parametrelili Kumaraswamy dağılımından gelen rasgele örneklem olmak üzere a ve b ye ait olabilirlik fonksiyonu,

$$L_{br\ddot{o}}(a, b; \mathcal{X}) = \prod_{i=1}^n f_{br\ddot{o}}(x_i; a, b) \quad (3)$$

$$= \prod_{i=1}^n f_{br\ddot{o}}(x_i; a, b)$$

$$= a^n b^n \prod_{i=1}^n x_i^{a-1} \prod_{i=1}^n (1 - x_i^a)^{b-1}$$

şeklinde yazılır.

Buradan logaritmik olabilirlik fonksiyonu,

$$\begin{aligned} \log L_{br\ddot{o}} &= \log \left(a^n b^n \prod_{i=1}^n x_i^{a-1} \prod_{i=1}^n (1 - x_i^a)^{b-1} \right) \\ &= \log a^n + \log b^n + \log \left(\prod_{i=1}^n x_i^{a-1} \right) \\ &\quad + \log \left(\prod_{i=1}^n (1 - x_i^a)^{b-1} \right) \\ &= n \log a + n \log b + \\ &\quad (a - 1) \sum_{i=1}^n \log(x_i) + (b - 1) \sum_{i=1}^n \log(1 - x_i^a) \end{aligned} \quad (4)$$

olarak elde edilir.

a ve b parametrelerinin basit rasgele örnekleme ile en çok olabilirlik tahmin edicilerini $\hat{a}_{br\ddot{o}}$ ve $\hat{b}_{br\ddot{o}}$ olarak gösterilsin. Yukarıda verilen olabilirlik fonksiyonunun sırasıyla a ve b parametrelerine göre türevleri alınır,

$$\frac{n}{a} + \sum_{i=1}^n \log(x_i) - (b - 1) \sum_{i=1}^n \frac{x_i^a \log a}{(1 - x_i^a)} = 0 \quad (5)$$

$$\frac{n}{b} + \sum_{i=1}^n \log(1 - x_i^a) = 0 \quad (6)$$

olarak bulunur ve (5) ile (6) eşitliklerinden,

$$\hat{b}_{br\ddot{o}} = \frac{-n}{\sum_{i=1}^n \log(1 - x_i^{\hat{a}_{br\ddot{o}}})} \quad (7)$$

ve $\hat{a}_{br\ddot{o}}$ nin lineer olmayan çözümü,

$$\begin{aligned} \frac{n}{\hat{a}_{br\ddot{o}}} + \sum_{i=1}^n \log(x_i) \\ - (b - 1) \sum_{i=1}^n \frac{x_i^{\hat{a}_{br\ddot{o}}} \log \hat{a}_{br\ddot{o}}}{(1 - x_i^{\hat{a}_{br\ddot{o}}})} = 0 \end{aligned} \quad (8)$$

olarak bulunur.

3.2. SKÖ için en çok olabilirlik tahmini

$X_1, X_2, \dots, X_n \sim F(x)$ dağılımlı bir rasgele örneklem olsun. Buradan elde edilen sıralı küme örneklemede $X_{(i:m)j}$: sıralı küme örneklemede r inci döngüdeki m birimlik örneklemede i inci sıradaki değişkeni temsil etmektedir. Burada $i = 1, \dots, m$ ve $j = 1, \dots, r$ olmak üzere örneklem büyüklüğü $n = m.r$ olan bir örneklem Kumaraswamy dağılımından sıralı

küme örnekleme ile elde edilsin. Gösterimi daha basit kılmak adına sıralı küme örnekleme rasgele değişkeni $X_{(i:m)j}$ yerine X_{ij} olarak gösterilecektir. Gözlemler sıralı küme örnekleme ile seçildiğinden dağılıma ait olasılık yoğunluk fonksiyonu yeniden ifade edilmelidir. X_{ij} rasgele değişkeninin olasılık yoğunluk fonksiyonu,

$$\begin{aligned} f_{sk\ddot{o}}(x_{ij}) &= \frac{m!}{(1-i)!(m-i)!} f(x_{ij}) [F(x_{ij})]^{i-1} \\ &\quad \times [1 - F(x_{ij})]^{m-i} \\ &= \frac{m!}{(1-i)!(m-i)!} a b x_{ij}^{a-1} \\ &\quad \times (1 - x_{ij}^a)^{b(m-i+1)-1} [1 - (1 - x_{ij}^a)^b]^{i-1} \end{aligned} \quad (9)$$

olarak yazılır. Burada $0 < x_{ij} < 1$ ve $a > 0, b > 0$ dir. a ve b nin olabilirlik fonksiyonu,

$$\begin{aligned} L_{sk\ddot{o}}(a, b; x) &= \prod_{j=1}^r \prod_{i=1}^m f_{rss}(x_{ij}; a, b) \\ &= \prod_{j=1}^r \prod_{i=1}^m \frac{m!}{(1-i)!(m-i)!} a b x_{ij}^{a-1} \\ &\quad \times \prod_{j=1}^r \prod_{i=1}^m (1 - x_{ij}^a)^{b(m-i+1)-1} [1 \\ &\quad - (1 - x_{ij}^a)^b]^{i-1} \end{aligned} \quad (10)$$

olur.

Buradan, logaritmik olabilirlik fonksiyonu,

$$\begin{aligned} \log L_{sk\ddot{o}}(a, b; x) &= m r \log(a) + m r \log(b) \\ &\quad + (a - 1) \sum_{j=1}^r \sum_{i=1}^m \log(x_{ij}) \\ &\quad + b \sum_{j=1}^r \sum_{i=1}^m (m - i + 1) \log(1 - x_{ij}^a) \\ &\quad - \sum_{j=1}^r \sum_{i=1}^m \log(1 - x_{ij}^a) \\ &\quad + \sum_{j=1}^r \sum_{i=1}^m (i - 1) \log[1 - (1 - x_{ij}^a)^b] \\ &\quad + r \sum_{i=1}^m \log \left(\frac{m!}{(i-1)!(m-i)!} \right) \end{aligned} \quad (11)$$

şeklinde yazılır.

a ve b nin sıralı küme örnekleme ile en çok olabilirlik tahmin edicileri $\hat{a}_{sk\ddot{o}}$ ve $\hat{b}_{sk\ddot{o}}$ gösterilsin. (12) nolu eşitlikteki sıralı küme örnekleme göre bulunan logaritmik olabilirlik fonksiyonunun a ve b ye göre türevleri sırasıyla (13) ve (14) nolu eşitliklerde gösterilmiştir. a' ya göre türev,

$$\begin{aligned}
& \frac{mr}{a} + \sum_{j=1}^r \sum_{i=1}^m \log(x_{ij}) \\
& -b \sum_{j=1}^r \sum_{i=1}^m (m-i+1) \frac{x_{ij}^a \log(x_{ij})}{1-x_{ij}^a} \\
& + \sum_{j=1}^r \sum_{i=1}^m \frac{x_{ij}^a \log(x_{ij})}{1-x_{ij}^a} \quad (12) \\
& + \sum_{j=1}^r \sum_{i=1}^m (i-1) \frac{bx_{ij}^a(1-x_{ij}^a)^{b-1} \log(x_{ij})}{1-(1-x_{ij}^a)^b} \\
& = 0
\end{aligned}$$

b' ye göre türev,

$$\begin{aligned}
& \frac{mr}{b} + \sum_{j=1}^r \sum_{i=1}^m (m-i+1) \log(1-x_{ij}^a) \\
& - \sum_{j=1}^r \sum_{i=1}^m (i-1) \frac{(1-x_{ij}^a)^b \log(1-x_{ij}^a)}{1-(1-x_{ij}^a)^b} = 0 \quad (13)
\end{aligned}$$

şeklinde elde edilir.

4. Genetik Algoritma

Genetik algoritmanın temelleri ilk olarak John Holland tarafından Michigian Üniversitesi'nde ortaya konulmuştur. 1975 yılında John Holland bu konudaki çalışmalarını "Adaptation in Natural and Artificial Systems" adlı kitabında toplamıştır [14]. Genetik algoritmalarındaki dikkat çeken ilerleme John Holland'ın öğrencisi David E. Goldberg tarafından 1985 yılında tamamlanan "Gaz Boru hatlarının Genetik Algoritma Kullanılarak Denetlenmesi" başlıklı doktora tezi ile gerçekleşmiştir. Bu çalışmasının ardından David E. Goldberg'ın 1989 yılında basılan "Makine Öğrenmesi, Arama ve Optimizasyon için Genetik Algoritma" adındaki kitabı genetik algoritmayı yeni bir seviyeye çıkarmıştır. Ayrıca bu kitap bugün dahi genetik algoritma konusunda en önemli referanslardan biridir [15].

Genetik algoritmalar doğadaki evrimsel süreci temel fikir olarak kabul eden optimizasyon problemlerinin çözümü için geliştirilmiş sezgisel bir eniyileme yöntemidir. Çok boyutlu arama uzayında doğal seçim ilkesine göre en güçlünün hayatta kaldığı bütünsel en iyi çözüm kümesini araştırır. Genetik algoritma rasgele olarak arama yöntemlerini esas alır parametre kodlama metotlarını kullanarak çalışır. Diğer optimizasyon metotlarına paralel olarak çözüm için tek bir yapının geliştirilmesinden ziyade bu tipte yapıların oluşturduğu bir küme geliştirir. Problemin mümkün olan çözümlerini ifade eden kümeye popülasyon adı verilir. Popülasyonu oluşturan bireylerin uygunluk fonksiyonundaki değerlerine göre optimal çözüme yakın uygunluk değeri olan bireylerin tekrar seçilme şansı yüksektir. Düşük uygunluk değerine sahip bireyler bir sonraki adım için seçilme ihtimalinin az olmasından dolayı belli

adımdan sonra tamamen nüfusun dışında kalırlar. Yeni popülasyon uygunluk değeri daha iyi olan bireylerden oluşur. Böylece her adımdaki popülasyon içi yüksek uygunluk değeri olan bireyler korunarak arama uzayında optimal bir çözüm bulunabilir.

Genetik algoritmalar diğer optimizasyon metotlarına göre son derece büyük arama uzayı olan problemlerde başarılı sonuçlar vermektedir. Genetik algoritma kullanılmasındaki esas hedef diğer analitik çözüm yollarının makul sonuçlar vermediği durumlarda optimizasyon problemlerine uygun çözümlerin aranmasında tercih edilir.

Genetik algoritma işleyişinde üç adet evrim operatörü mevcuttur. Bu operatörler seçim, çaprazlama ve mutasyon operatörleri olarak ifade edilir. Bu operatörler genetik algoritma içindeki süreçte bir sonraki adımda oluşacak olan popülasyondaki her bir birey için uygulanır.

4.1. Neden genetik algoritma?

Aslında Kumaraswamy dağılımının a ve b parametrelerinin en çok olabilirlik tahmininin hesaplanmasında olabilirlik fonksiyonunu maksimum yapan a ve b değerlerinin bulunması istenmektedir. Buna bağlı olarak maksimize edilmek istenen olabilirlik fonksiyonu bir kısıtlı optimizasyon problemi olarak yazılabilir. Burada kısıtlar parametrelerin pozitif olmasıdır.

$$\begin{aligned}
& \max L(a, b; x) \\
& a > 0 \\
& b > 0 \quad (14)
\end{aligned}$$

Burada ele alınan, olabilirlik fonksiyonunu maksimum yapan pozitif a ve b değerlerinin bulunması problemindeki gösterim literatürdeki genel kısıtlı optimizasyon gösterimine uygun olarak aşağıdaki gibi yazılır.

$$\begin{aligned}
& \min -L(a, b; x) \\
& -a < 0 \\
& -b < 0 \quad (15)
\end{aligned}$$

Bu türdeki kısıtlı optimizasyon problemlerinde Karush-Kuhn-Tucker koşulları ile çözüm araştırılabilir.

f minimumu aranan fonksiyon ve $\mu_i, i = 1, 2, \dots, m$ Karush-Kuhn-Tucker çarpanları olmak üzere, fonksiyonu minimum yapan noktanın bulunması için Karush-Kuhn-Tucker koşulları,

$$\begin{aligned}
& i) \mu \geq 0 \\
& ii) \nabla f(x) + \mu^T = 0 \\
& iii) \mu g(x) = 0 \quad (16)
\end{aligned}$$

şeklinde dir.

Bu koşulları sağlayan Lagrange Algoritmasından faydalanılarak bu tür problemlerin çözümü mümkün olabilir [16]. Ayrıca bu yönteminin dışında klasik bir yaklaşım olan Newton-Raphson yöntemi de tercih edilebilir. Ancak bu yollar ile bu problem için çözüm denendiğinde yakınsama ve başlangıç değerlerin belirlenmesi gibi birtakım problemlerle karşılaşmıştır. Genetik algoritma ise incelenen fonksiyonun türev bilgisine ihtiyaç duymadığından ve hesaplamalarda kısıtların dikkate alınmasını kolaylaştırdığından bu çalışmada parametrelerin tahmin edilmesinde Genetik Algoritmaların uygulanması araştırılmıştır.

5. Simülasyon Çalışması

Burada Kumaraswamy dağılımı parametreleri için olabilirlik fonksiyonlarının genetik algoritma ile maksimum noktalarının araştırılmasında R yazılımı kullanılmıştır [17]. R yazılımında kullanıcılara büyük kolaylık sağlayan başka araştırmacılar tarafından hazırlanmış birçok konuya ilişkin paketler bulunmaktadır. Bu simülasyon çalışmasında R da çalışma prensibi en çok olabilirlik yöntemine dayanan "fitdistrplus" paketi ve Newton-Raphson yöntemine göre hesaplamalar yapan "optimx" paketleri denenmiş ancak her iki pakette de yakınsama sorunu ile karşılaşmıştır [18] [19]. Genetik algoritma ile ilgili hesaplamalar için "GA" paketi kullanılmıştır [20]. Ayrıca Kumaraswamy

dağılımından sayı üretmek için "VGAM" paketi kullanılmıştır [21].

Kumaraswamy dağılımının bilinmeyen parametrelerine ait basit rasgele örneklem ve sıralı küme örneklemesi için en çok olabilirlik tahmin edicilerinin performanslarını karşılaştırabilmek için yan, hata kareler ortalaması ve etkinlikler hesaplanmıştır. Örneklem çapı 20, 30 ve 60 seçilmiş ve her bir deney için tekrar sayısı 1000 olarak belirlenmiştir. Tahmin edicilerin kıyaslanmasında yan, hata kareler ortalaması ve etkinlik aşağıdaki eşitliklerdeki gibi ifade edilmektedir.

$$Yan(\hat{a}) = E[\hat{a}] - a \quad (17)$$

$$HKO(\hat{a}) = E_{\hat{a}}[(\hat{a} - a)^2] = Var(\hat{a}) + Yan(\hat{a})^2 \quad (18)$$

$\hat{a}_{brö}$, a parametresi için basit rasgele örneklem ile elde edilen tahmin edici ve $\hat{a}_{skö}$, sıralı küme örneklemesi ile elde edilen tahmin ediciyi göstermektedir. Buna göre etkinlik (20) nolu eşitlikteki gibidir.

$$Etk(\hat{a}_{skö}) = \frac{HKO(\hat{a}_{brö})}{HKO(\hat{a}_{skö})} \quad (19)$$

Simülasyon çalışmasına ait sonuçlar Tablo 1-6' da sunulmuştur.

Tablo 1. $a=0.5$ ve $n=20$ için Yan, HKO ve Etkinlik

b	n	Yan ($\hat{a}_{brö}$)	HKO ($\hat{a}_{brö}$)	Yan ($\hat{b}_{brö}$)	HKO ($\hat{b}_{brö}$)	m	r	Yan ($\hat{a}_{skö}$)	HKO ($\hat{a}_{skö}$)	etk ($\hat{a}_{skö}$)	Yan ($\hat{b}_{brö}$)	HKO ($\hat{b}_{brö}$)	etk ($\hat{b}_{brö}$)	
0,5	20	0,0884	0,0602	0,0574	0,0318	2	10	0,0753	0,0543	1,1083	0,0485	0,0224	1,4192	
							4	5	0,0597	0,0325	1,8500	0,0423	0,0166	1,9030
							5	4	0,0512	0,0296	2,0312	0,0361	0,0162	1,9607
1	20	0,0654	0,0325	0,1673	0,1764	2	10	0,0508	0,0253	1,2811	0,1230	0,1374	1,2839	
							4	5	0,0458	0,0204	1,5865	0,1124	0,1119	1,5762
							5	4	0,0418	0,0177	1,8283	0,0978	0,0861	2,0485
3	20	0,0255	0,0129	0,4361	1,2569	2	10	0,0254	0,0112	1,1424	0,4012	1,1756	1,0690	
							4	5	0,0246	0,0080	1,6030	0,3468	0,9529	1,3189
							5	4	0,0251	0,0080	1,6128	0,3253	0,8979	1,3996

Tablo 2. $a=0.5$ ve $n=30$ için Yan, HKO ve Etkinlik

b	n	Yan ($\hat{a}_{brö}$)	HKO ($\hat{a}_{brö}$)	Yan ($\hat{b}_{brö}$)	HKO ($\hat{b}_{brö}$)	m	r	Yan ($\hat{a}_{skö}$)	HKO ($\hat{a}_{skö}$)	etk ($\hat{a}_{skö}$)	Yan ($\hat{b}_{brö}$)	HKO ($\hat{b}_{brö}$)	etk ($\hat{b}_{brö}$)	
0,5	30	0,0619	0,0359	0,0379	0,0169	2	15	0,0540	0,0270	1,3293	0,0373	0,0142	1,1920	
							3	10	0,0371	0,0210	1,7068	0,0239	0,0107	1,5721
							5	6	0,0394	0,0173	2,0798	0,0250	0,0083	2,0239
1	30	0,0365	0,0179	0,0901	0,0939	2	15	0,0318	0,0162	1,1007	0,0876	0,0849	1,1050	
							3	10	0,0308	0,0127	1,4065	0,0812	0,0699	1,3422
							5	6	0,0225	0,0106	1,6797	0,0553	0,0488	1,9225
3	30	0,0246	0,0094	0,3250	0,9495	2	15	0,0202	0,0077	1,2232	0,2761	0,8086	1,1741	
							3	10	0,0160	0,0066	1,4278	0,2517	0,7559	1,2560
							5	6	0,0164	0,0051	1,8465	0,2522	0,6400	1,4836

Tablo 3. a=0.5 ve n=60 için Yan, HKO ve Etkinlik

b	n	Yan ($\hat{a}_{br\bar{o}}$)	HKO ($\hat{a}_{br\bar{o}}$)	Yan ($\hat{b}_{br\bar{o}}$)	HKO ($\hat{b}_{br\bar{o}}$)	m	r	Yan ($\hat{a}_{sk\bar{o}}$)	HKO ($\hat{a}_{sk\bar{o}}$)	etk ($\hat{a}_{sk\bar{o}}$)	Yan ($\hat{b}_{br\bar{o}}$)	HKO ($\hat{b}_{br\bar{o}}$)	etk ($\hat{b}_{br\bar{o}}$)
0,5	60	0,0326	0,0138	0,0201	0,0071	2	30	0,0239	0,0120	1,1473	0,0172	0,0063	1,1193
							20	0,0241	0,0094	1,4606	0,0148	0,0044	1,6079
							15	0,0207	0,0083	1,6665	0,0142	0,0040	1,7622
1	60	0,0206	0,0083	0,0433	0,0331	2	30	0,0200	0,0071	1,1664	0,0371	0,0298	1,1116
							20	0,0178	0,0058	1,4185	0,0416	0,0267	1,2382
							15	0,0145	0,0048	1,7048	0,0349	0,0231	1,4288
3	60	0,0140	0,0049	0,1901	0,4961	2	30	0,0117	0,0036	1,3360	0,1847	0,4285	1,1578
							20	0,0104	0,0033	1,4545	0,1563	0,3923	1,2645
							15	0,0090	0,0030	1,6045	0,1368	0,3455	1,4357

Tablo 4. a=2 ve n=20 için Yan, HKO ve Etkinlik

b	n	Yan ($\hat{a}_{br\bar{o}}$)	HKO ($\hat{a}_{br\bar{o}}$)	Yan ($\hat{b}_{br\bar{o}}$)	HKO ($\hat{b}_{br\bar{o}}$)	m	r	Yan ($\hat{a}_{sk\bar{o}}$)	HKO ($\hat{a}_{sk\bar{o}}$)	etk ($\hat{a}_{sk\bar{o}}$)	Yan ($\hat{b}_{br\bar{o}}$)	HKO ($\hat{b}_{br\bar{o}}$)	etk ($\hat{b}_{br\bar{o}}$)
0,5	20	0,3820	0,9413	0,0609	0,0295	2	10	0,29071	0,74390	1,26540	0,04892	0,02308	1,2799
							5	0,2455	0,5520	1,7052	0,0381	0,0169	1,7451
							4	0,2286	0,5234	1,7984	0,0340	0,0147	2,0056
1	20	0,2465	0,5367	0,1665	0,1923	2	10	0,2272	0,4148	1,2938	0,1351	0,1375	1,3984
							5	0,1772	0,3358	1,5981	0,1220	0,1038	1,8523
							4	0,1742	0,3034	1,7687	0,1098	0,0956	2,0102
3	20	0,1387	0,2282	0,4324	1,3410	2	10	0,1113	0,1743	1,3090	0,4367	1,2156	1,1031
							5	0,0757	0,1365	1,6714	0,3241	0,9791	1,3696
							4	0,1068	0,1368	1,6679	0,3727	0,9896	1,3549

Tablo 5. a=2 ve n=30 için Yan, HKO ve Etkinlik

b	n	Yan ($\hat{a}_{br\bar{o}}$)	HKO ($\hat{a}_{br\bar{o}}$)	Yan ($\hat{b}_{br\bar{o}}$)	HKO ($\hat{b}_{br\bar{o}}$)	m	r	Yan ($\hat{a}_{sk\bar{o}}$)	HKO ($\hat{a}_{sk\bar{o}}$)	etk ($\hat{a}_{sk\bar{o}}$)	Yan ($\hat{b}_{br\bar{o}}$)	HKO ($\hat{b}_{br\bar{o}}$)	etk ($\hat{b}_{br\bar{o}}$)
0,5	30	0,2613	0,5165	0,0463	0,0172	2	15	0,1904	0,4502	1,1472	0,0353	0,0137	1,2554
							10	0,1606	0,3516	1,4690	0,0260	0,0108	1,5944
							6	0,1281	0,2695	1,9160	0,0226	0,0081	2,1260
1	30	0,1243	0,2881	0,0874	0,0894	2	15	0,1165	0,2293	1,2562	0,0809	0,0737	1,2129
							10	0,1423	0,2276	1,2656	0,0817	0,0672	1,3304
							6	0,1190	0,1585	1,8177	0,0731	0,0492	1,8151
3	30	0,1077	0,1438	0,3260	0,8816	2	15	0,0892	0,1224	1,1744	0,3065	0,8250	1,0685
							10	0,0748	0,1113	1,2917	0,2643	0,7618	1,1572
							6	0,0656	0,0835	1,7221	0,2527	0,6235	1,4139

Tablo 6. a=2 ve n=60 için Yan, HKO ve Etkinlik

b	n	Yan ($\hat{a}_{br\bar{o}}$)	HKO ($\hat{a}_{br\bar{o}}$)	Yan ($\hat{b}_{br\bar{o}}$)	HKO ($\hat{b}_{br\bar{o}}$)	m	r	Yan ($\hat{a}_{sk\bar{o}}$)	HKO ($\hat{a}_{sk\bar{o}}$)	etk ($\hat{a}_{sk\bar{o}}$)	Yan ($\hat{b}_{br\bar{o}}$)	HKO ($\hat{b}_{br\bar{o}}$)	etk ($\hat{b}_{br\bar{o}}$)
0,5	60	0,1026	0,2217	0,0211	0,0072	2	30	0,0917	0,1686	1,3151	0,0133	0,0049	1,4573
							20	0,0836	0,1588	1,3960	0,0129	0,0048	1,4972
							15	0,0908	0,1484	1,4943	0,0161	0,0044	1,6332
1	60	0,1001	0,1517	0,0540	0,0361	2	30	0,0739	0,1118	1,3567	0,0404	0,0288	1,2504
							20	0,0447	0,0878	1,7273	0,0271	0,0222	1,6234
							15	0,0487	0,0791	1,9177	0,0267	0,0220	1,6408
3	60	0,0557	0,0775	0,2058	0,5260	2	30	0,0334	0,0642	1,2081	0,1510	0,4397	1,1964
							20	0,0356	0,0540	1,4356	0,1552	0,4038	1,3027
							15	0,0510	0,0515	1,5064	0,1721	0,3769	1,3956

6. Tartışma ve Sonuç

Burada en çok olabilirlik yöntemi ile parametre tahmin edilirken kolay bir şekilde türevi alınamayan karmaşık olabilirlik fonksiyonların

çözümünde ve bu problemde olduğu gibi parametrelere ilişkin kısıtların olduğu durumlarda çözüm için genetik algoritmanın değerlendirilebileceği gösterilmiştir. Özellikle sıralı küme örneklemesinde en çok olabilirlik tahmin

edicisinin bulunması problemi kısıtlı bir optimizasyon problemi olarak ifade edilebilir. Bu problemin çözümü için Newton-Raphson veya Karush-Kuhn-Tucker koşullarını temel alan Lagrange algoritması düşünülebilir. Ancak bu problemlerin çözümünde yakınsama ve başlangıç değerlerinin seçimine bağlı olarak hesaplama problemleri olduğu saptanmıştır. Bu nedenle bu tür durumlarda genetik algoritma yaklaşımının iyi bir alternatif olabileceği görülmüştür.

Simülasyon sonuçlarında görüldüğü üzere sıralı küme örneklemesi ile elde edilen tahmin edicilerin basit rasgele örnekleme ile elde edilen tahmin edicilere kıyasla daha küçük yan ve hata kareler ortalamasına sahip olduğu ve hata kareler ortalaması oranlarıyla bulunan etkinlikler açısından da daha iyi olduğu gösterilmiştir.

Sonuç olarak, bu çalışmada olduğu gibi, parametrelere ait kısıtların olduğu tahmin problemlerinde analitik metotlarının sonuç vermediği ya da uygulanmasının problemleri olduğu durumlarda sezgisel bir yaklaşım olarak genetik algoritma tercih edilebilir.

Kaynakça

- [1] Kumaraswamy, P. 1980. A generalized Probability Density Function for Double-Bounded Random Processes. *Journal of Hydrology*, 46(1980), 79-88.
- [2] Jones, M. C. 2009. Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, 6(2009), 70-81.
- [3] Hussian, M. A. 2014. Bayesian and Maximum Likelihood Estimation for Kumaraswamy Distribution based on Ranked Set Sampling. *American Journal of Mathematics and Statistics*, 4(2014), 30-37.
- [4] McIntyre, G. A. 1952. A Method for Unbiased Selective Sampling, using Ranked Sets. *Australian Journal of Agricultural Research*, 1952 385-390.
- [5] Patil, G. P., Surucu, B. and Egemen D. 2002. Ranked set sampling. *Wiley StatsRef: Statistics Reference Online*, 2002.
- [6] Takahasi, K., and Wakimoto, K. 1968. On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics*, 1968, 1-31.
- [7] Dell, T. R., and Clutter, J. L. 1972. Ranked set sampling theory with order statistics background. *Biometrics*, 1972, 545-555.
- [8] Stokes, S. L. 1977. Ranked Set Sampling with Concomitant Variables. *Communications in Statistics-Theory and Methods*, 1977, 1207-1211.
- [9] Samawi, H. M. 1996. Stratified Ranked Set Sample. *Pakistan Journal of Statistics-All Series*, 12(1996), 9-16.
- [10] Samawi, H. M., Ahmed, M. S., and Abu-Dayyeh, W. 1996. Estimating the Population Mean using Extreme Ranked Set Sampling. *Biometrical Journal*, 38(1996), 577-586.
- [11] Al-Saleh, M. F., and Al-Kadiri M. A. 2000. Double-Ranked Set Sampling. *Statistics & Probability Letters*, 48(2000), 205-212.
- [12] Al-Saleh, M. F., and Al-Omari A. I. 2002. Multistage Ranked Set Sampling. *Journal of Statistical planning and Inference*, 102(2002), 273-286.
- [13] Muttlak, H. A. 2003. Investigating the Use of Quartile Ranked Set Samples for Estimating the Population Mean. *Applied Mathematics and Computation*, 146(2003), 437-443.
- [14] Holland, J.H. 1975. *Adaptation in Natural and Artificial Systems*. MIT Press.
- [15] Goldberg, D. 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley.
- [16] Chong, E. KP., Zak, S. H. 2013. *An Introduction to Optimization*. 2nd, John Wiley & Sons.
- [17] R Core Team (2017). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria., [Çevrimiçi]. Available: <https://www.R-project.org/>.
- [18] C. D. Marie Laure Delignette-Muller, *fitdistrplus: An R Package for Fitting Distributions*, *Journal of Statistical Software* 64(4), 1-34, 2015. [Çevrimiçi]. Available: <http://www.jstatsoft.org/v64/i04/>.
- [19] R. V. John C. Nash, *Unifying Optimization Algorithms to Aid Software System Users: optimx for R*, *Journal of Statistical Software*, 43(9), 1-14, 2011. [Çevrimiçi]. Available: <http://www.jstatsoft.org/v43/i09/>.
- [20] Scrucca, L., GA: A Package for Genetic Algorithms in R., *Journal of Statistical Software*, 53(4), 1-37., 2013. [Çevrimiçi]. Available: <http://www.jstatsoft.org/v53/i04/>.
- [21] Yee, T. W., *The VGAM Package for Categorical Data Analysis*, *Journal of Statistical Software*, 32(10), 1-34., 2010. [Çevrimiçi]. Available: <http://www.jstatsoft.org/v32/i10/>.